Journal of NeuroEngineering and Rehabilitation

RESEARCH

Open Access

# Validity and usability for digital cognitive assessment tools to screen for mild cognitive impairment: a randomized crossover trial

Min Wu[1], Jialin Feng[1], Ruini Sun[1], Shangyang Zhang[2], Yongjin Zhang[1], Fengjie Yang[3], Xiaoyi Zhang[4], Yunxian Ye[4], Ni Gong[2] and Jing Liao[1,5,6*]

## Abstract

**Background** The practicality of implementing digital cognitive screening tests in primary health care (PHC) for the detection of cognitive impairments, particularly among populations with lower education levels, remains unclear. The aim of this study is to assess the validity and usability of digital cognitive screening tests in PHC settings.

**Methods** We utilized a randomized crossover design, whereby 47 community-dwelling participants aged 65 and above were randomized into two groups. One group completed the paper-based Mini-Mental State Examination (MMSE) and Clock Drawing Test (CDT) first, followed by the tablet-based digital version after a two-week washout period, while the other group did the reverse. Validity was assessed by Spearman correlation, linear mixed-effects models, sensitivity specificity, and area under the curve (AUC). Usability was assessed through the Usefulness, Satisfaction, and Ease of Use (USE) questionnaire, participant preferences and assessment duration. Regression analyses were conducted to explore the impact of usability on digital test scores, controlling for cognitive level, education, age, and gender.

**Results** Regarding validity, digital tests showed moderate correlations with paper-based versions and superior AUC performance. The AUC was 0.65 for the MMSE versus 0.82 for the electronic MMSE (eMMSE), and 0.45 for the CDT compared to 0.65 for the electronic CDT (eCDT). Regarding usability, while older participants gave positive feedback on digital tests ($P < 0.001$), they preferred paper-based versions. The eMMSE took significantly longer to complete than the MMSE, averaging 7.11 min versus 6.21 min ($P = 0.01$). Notably, digital test scores were minimally affected by subjective attitudes but strongly linked to test duration ($\beta = -0.62$, *95% CI*: -1.07 to -0.17).

**Conclusions** Digital cognitive tests are valid and feasible in PHC settings but face implementation challenges, especially in usability and adaptability among individuals with lower education levels.

**Keywords** Mild cognitive impairment, Digital cognitive assessment, Validity, Usability

*Correspondence:
Jing Liao
liaojing5@mail.sysu.edu.cn
[1]Department of Medical Statistics and Epidemiology, Sun Yat-sen University, Guangzhou, Guangdong, China
[2]School of Nursing, Jinan University, Guangzhou, Guangdong, China
[3]Qiaonan Street Community Health Service Centre, Panyu District, Guangzhou, China
[4]Luopu Street Community Health Service Centre, Panyu District, Guangzhou, China
[5]Global Health Institute, School of Public Health, Institute of State Governance, Sun Yat-sen University, Guangzhou, China
[6]Department of Medical Statistics and Epidemiology, School of Public Health, Institute of State Governance, Sun Yat-sen University, No.74 Zhongshan 2nd Road, Guangzhou 51008, China

## Background

Chinese dementia patients account for quarter of the global total, with annual medical and care costs reaching 1.06 trillion yuan, which is 1.3 times the global average [1]. With no cost-effective drugs for dementia, early screening and intervention for mild cognitive impairment (MCI), the transitional stage between normal cognition and dementia, are crucial for prevention and treatment [2]. China recently piloted primary healthcare (PHC)-based dementia screening for older adults aged 65 years and above in several cities, while the rate of missed diagnoses remains high, with less than 10% of dementia cases detected during screening [3].

The lower detection rate may result from lack of easy-to-implement screening tools among other reasons. The current widely-used Mini-Mental State Examination (MMSE) requires professional training for the examiner [3]. The manual scoring, recording, and analysis are cumbersome and subject to examiner bias, consuming significant time and human resources [3]. On the other hand, standardized and scalable digital screening tools may be better suited for early MCI screening in PHC [4]. Standardized procedures and automated scoring systems reduce reliance on professionals and simplify data management [5], thereby may enhance screening accuracy and efficiency.

Digital cognitive screening tests hold great promise as tools in PHC settings, but their effectiveness remains insufficiently addressed. A recent systematic review revealed that most validation studies on digital cognitive tests are concentrated in Western countries (13/20), with limited evidence from China [6]. Only a few studies have explored the development and application of digital cognitive tests in China, including efforts in Hong Kong to develop new digital cognitive tests [7, 8] and in mainland China to digitize the Beijing version of the MoCA [9]. However, the validity results of these studies show notable variation. The correlation between digital and traditional paper-based tests ranges from 0.67 to 0.93, while the AUC for distinguishing MCI from cognitively normal individuals ranges from 0.78 to 0.97 [7–9]. This variation may be attributed to differences in the educational levels of study populations. For instance, the digital MoCA demonstrated the highest correlation (0.93) [9] in populations with an average of 15 years of education, whereas study with 46.9% of participants having six or fewer years of education reported a lower correlation (0.71) [8]. These findings suggest that while preliminary evidence supports the potential of digital cognitive tests, further research is needed to validate their practicality across populations with varying educational backgrounds. Given that the MMSE and CDT scales are less influenced by educational attainment [10, 11] and are widely used in community screening in China [12], adapting them into digital formats and validating their effectiveness could significantly enhance the practicality of digital cognitive tests.

Moreover, unlike paper-based tests, digital cognitive screening tests as software programs, yet existing studies often overlook the impact of usability on their overall effectiveness. A recent systematic review found that only two studies employed usability questionnaires [13, 14] to confirm the usability of digital cognitive tests. Given that digital tests may be influenced by digital literacy and technological adaptability, and that digital literacy is generally lower among older adults with low educational levels. Wallace. S [15] demonstrated that greater experience with touchscreen devices, such as smartphones and tablets, was associated with higher scores on the electronic MoCA. Currently, only one study by Lunardini. F [14] evaluated dimensions such as intention to use, perceived usefulness, ease of learning, and actual use among 83 older adults, finding high technological adaptability. However, this study was descriptive in nature and did not assess the impact of these dimensions on the performance of digital cognitive tests. Notably, the study populations were primarily composed of highly educated individuals from Western countries, limiting the generalizability of the findings to Asian populations. Therefore, it is essential to evaluate the system usability of digital cognitive tests among Chinese populations and its impact on test performance.

Our study aims to evaluate the validity of digitally adapted cognitive screening tests in PHC settings and their usability across different education levels, as well as to explore the impact of usability on the performance of digital cognitive screening tests. First-hand evidence will be provided for the practical application of digital screening tests in PHC settings.

## Methods

### Study design

The study was conducted in a PHC covering both urban and rural areas in Guangzhou, China, from November 2023 to January 2024. A randomized cross-over design was adopted (Fig. 1). Eligible participants were randomly assigned to one of two groups at 1:1 ratio: one group completed the paper-based cognitive assessment first, followed by the digital version after a two-week washout period, while the other group did the reverse. Participants with a positive result on either MMSE test and a randomly selected 10% sample of those tested negative, were verified by professional geriatric neurologists as normal, MCI, or dementia according to the International Classification of Diseases, 11th Revision (ICD-11) [16] and Peterson's criteria [17]. Ethic consents were obtained from the Biomedical Research Ethics Committee of the School of Public Health, Sun Yat-sen University, China

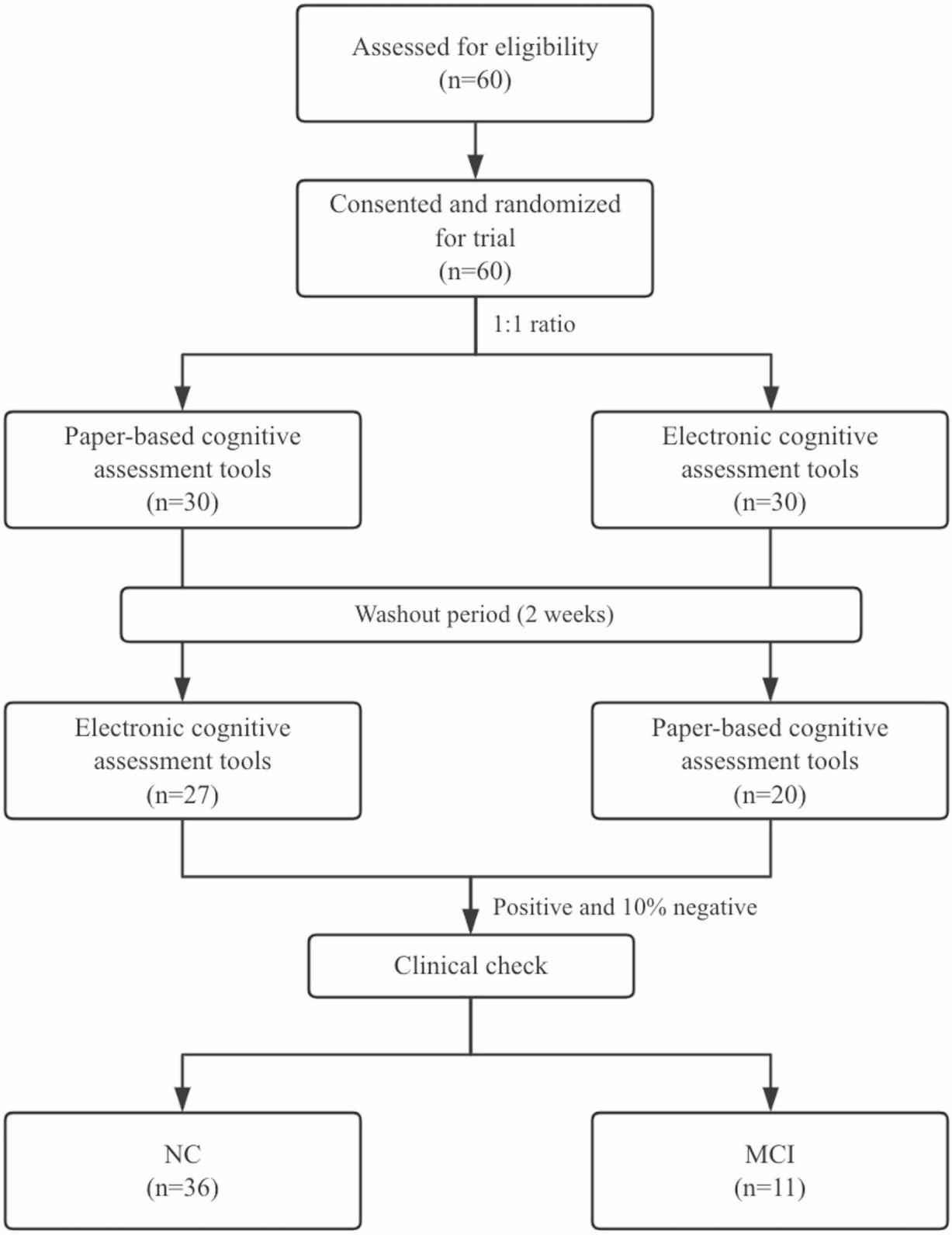


**Fig. 1** Participant flowchart
Notes: MCI, Mild cognitive impairment; NC, Normal cognition

(ethics approval number: 2023 – 129). This study was conducted in accordance with the ethical standards set forth in the 1964 Declaration of Helsinki and its subsequent amendments. All subjects provided written informed consent voluntarily.

### Study participants

Participants aged 65 years and over, were recruited through chronic disease management programs and health service announcements via posters, WeChat messages, and phone invitations. Eligibility criteria were (1) residents in the area (residing for over 6 months), (2) capable to read enlarged printed text, see a screen, and (3) consent to participate. Exclusion criteria were participants with (1) other neurological or psychiatric disorders affecting cognition, (2) significant sensory impairments, and (3) were taking medication that may affect cognitive function, such as anticholinergic drugs.

### Sample size

Sample size was calculated according to the correlation between the digital version and the paper version, specifically between the eMMSE and paper-based MMSE, and between the eCDT and paper-based CDT. Previous validation studies on digital adaptations of these tools such as the MoCA and the Self-Administered Gerocognitive Examination (SAGE) have reported correlation coefficients ranging from 0.68 to 0.93 [9, 15, 18]. Based on these findings, we conservatively assumed a moderate correlation coefficient ($\rho = 0.6$), with a two-sided alpha of 0.05 and statistical power of 0.90. Using PASS 2021 software, we estimated that a minimum of 38 participants would be required to detect this correlation, after accounting for an anticipated 20% dropout rate. To further ensure data robustness, we conservatively set a higher recruitment target of 60 participants. Ultimately, 47 participants completed both the e-version and paper version of the assessments, yielding a follow-up rate of 78.3%.

### Randomization

After enrollment, participants were matched based on age, gender, and educational level, and then randomly assigned in a 1:1 ratio to receive either the digital-first or paper-based-first version of the cognitive assessments. Randomization sequences were generated using R software (version 4.2.2).

### Paper-based and digital cognitive tests

#### *MMSE*

The paper-based MMSE includes 30 individual tests covering five main dimensions: orientation, registration, attention and calculation, recall, and language [19]. Scores below 27 for participants with ≥6 years of education, 23 for those with 1 to 5 years, and 20 for illiterate individuals [20] are considered positive for screening.

#### *CDT*

The CDT is a command-directed clock drawing test used to assess executive functions [21]. In the paper-based version, participants are instructed to "draw a clock face, including numbers and hands, setting the hands to 11:10, so that a child can understand it." The total score is 15, with a cut-off value of 10 points [22].

#### *eMMSE and eCDT*

The electronic versions of the MMSE (eMMSE) and the CDT (eCDT) were administered on tablet devices, with the same questions and rating criteria as their paper-based versions. All test instructions were presented to participants through pre-recorded standardized audio prompts. For most verbal response items, participants answered orally, while PHC healthcare providers recorded and scored the responses in real time on a separate synchronized tablet according to standard criteria. For drawing tasks involving executive function and visuospatial skills, such as clock drawing in the eCDT and the figure-copying task in the eMMSE, participants used their fingers to draw directly on the tablet screen. The system itself did not feature automated scoring.

### Outcome measures

#### *Validity*

Our study mainly examined the criterion validity of the digital cognitive tests, which was evaluated by comparing their total and single-dimension scores with the paper-based versions in terms of correlation and differences. Additionally, we compared their sensitivity, specificity, Youden's index and the area under the curve (AUC) using neurologist-verified results as the gold standard.

#### *Usability*

The Usefulness, Satisfaction, and Ease of Use (USE) questionnaire [23] was administered after completion of the digital tests in order to assess participants' attitudes towards the usability. This questionnaire contains 30 items on a 7-point Likert scale from strongly disagree (1) to strongly agree (7), which measures across four dimensions: usefulness, ease of use, ease of learning and satisfaction. Participants' preferences were also recorded (1 = paper-based; 2 = digital; 3 = unsure/no preference), with open-ended question collecting their evaluations of both versions, and their suggestions for future improvements.

Further, we recorded the time each participant took to complete both versions of the cognitive tests. Times were recorded manually for the paper versions, and automatically for the digital versions. It should be noted that we

**Table 1** Demographic characteristics and cognitive scores of study participants ($n=47$)

| Variable | |
|---|---|
| **Age in years** | 72.65 ± 6.09 |
| **Education level** | |
| Primary (≤ 6 years) | 19.14% |
| Middle (7–12 years) | 40.43% |
| High (>12 years) | 40.43% |
| **Sex (Female)** | 68.09% |
| **Total MMSE score** | 26.72 ± 3.00 |
| Orientation | 9.45 ± 1.08 |
| Registration | 2.83 ± 0.38 |
| Attention and calculation | 4.02 ± 1.15 |
| Recall | 2.51 ± 0.86 |
| Language | 7.89 ± 1.09 |
| **Total eMMSE score** | 25.87 ± 3.29 |
| Orientation | 9.49 ± 1.08 |
| Registration | 2.53 ± 0.91 |
| Attention and calculation | 4.06 ± 1.28 |
| Recall | 2.04 ± 1.04 |
| Language | 7.72 ± 1.21 |
| **Total CDT score ($n=43$)** | 10.79 ± 3.78 |
| **Total eCDT score ($n=43$)** | 10.78 ± 3.08 |

Notes: Values are shown as Mean ± SD (range) or %. MMSE, Mini-Mental State Examination; CDT, Clock Drawing Test; eMMSE, electronic versions of the Mini-Mental State Examination; eCDT, electronic versions of Clock Drawing Test; SD, Standard Deviation

did not account for the time required to input paper-based data into the electronic record system.

### Statistical analyses

Spearman correlation coefficients assessed the correlation between the paper-based and digital versions. Linear mixed-effects models were used to estimate the scores of the tools, incorporating fixed effects to test for differences between the paper-based and digital versions, and random effects for participants to account for repeated measures and individual variability. Sensitivity, specificity, Youden's index and AUC were obtained.

One-sample Wilcoxon signed rank test was used to compare the scores of USE question with the value 4 (the median of the Likert scale), and paired t-test for comparing the time taken to complete the tools in paper-based and digital versions. Preferences were described by component ratios (%) and open-ended questions were summarized in excel to explain the reasons for the preferences. ANOVA or chi-square tests were used to compare differences in usability across education levels.

Regression analyses were performed to control for the effects of demographic variables and usability on digital cognitive tests scores. Interaction terms (e.g., education level * duration and education level * USE) were included to examine potential moderating effects but were found to be non-significant. For simplicity, the final model did not include these interaction terms. Robust regression procedures were used in order to accommodate the skewed distribution of scores and to reduce the weight of any extreme outliers.

All statistical analyses were conducted using R (version 2023), with *P*-values less than or equal to 0.05 were considered statistically-significant.

**Table 2** Correlation and differences between paper-based and digital cognitive tests

| Variable (paper vs. digital) | Correlation | | Differences | |
|---|---|---|---|---|
| | Coefficient | *P* | Estimate | *P* |
| **Total MMSE score ($n=47$)** | 0.59 | < 0.001 | -0.79 | 0.02 |
| Orientation | 0.49 | < 0.01 | 0.05 | 0.76 |
| Time-orientation | 0.37 | 0.01 | 0.094 | 0.39 |
| Space-orientation | 0.37 | 0.01 | -0.044 | 0.70 |
| Registration | -0.05 | 0.72 | -0.29 | 0.05 |
| Attention and calculation | 0.56 | < 0.001 | 0.05 | 0.77 |
| Recall | 0.15 | 0.31 | -0.41 | 0.02 |
| Language | 0.28 | 0.06 | -0.18 | 0.40 |
| Naming | / | / | -0.04 | 0.16 |
| Repetition | 0.00 | 0.98 | -0.04 | 0.62 |
| Comprehension | -0.08 | 0.59 | 0.09 | 0.14 |
| Following commands | -0.03 | 0.85 | -0.08 | 0.60 |
| Visuospatial skill | 0.17 | 0.26 | -0.11 | 0.18 |
| **CDT ($n=43$)** | 0.62 | < 0.001 | 0.06 | 0.89 |

Notes: MMSE, Mini-Mental State Examination; CDT, Clock Drawing Test; SD, Standard Deviation. Correlations were assessed by Spearman's correlation and differences were analyzed by linear mixed effects

## Results

The demographic characteristics and cognitive scores are displayed in Table 1. Nearly half of the participants were over 70 years old, with a male-to-female ratio of 1:1.5, and a middle-level education or higher. The mean total MMSE score was 26.72 ± 3.00, while the eMMSE score was 25.87 ± 3.29. The CDT and eCDT scores were averaging around 10.79 and 10.78, respectively.

### Validity of digital cognitive tests

The Spearman correlation between MMSE and eMMSE total scores was 0.59 ($P<0.001$) (Table 2). The linear mixed-effect model indicated that the global MMSE score was 0.79 points higher than the total eMMSE score (*95% CI* = 0.14 to 1.43, $P=0.02$). The correlation between CDT and eCDT scores was 0.62 ($P<0.001$). There were no statistically significant differences in the total scores ($P=0.89$). The sensitivity of the MMSE was 0.55, compared to 0.91 for the eMMSE, while the specificity was 0.75 for both versions. The AUC was 0.65 (*95% CI* = 0.42 to 0.88) for the MMSE and 0.82 (*95% CI* = 0.69 to 0.94) for the eMMSE. For the CDT, the AUC was 0.45 (*95% CI* = 0.24 to 0.67) compared to 0.65 (*95% CI* = 0.44 to 0.85) for the eCDT (Table 3).

**Table 3** Sensitivity and specificity of paper-based and digital cognitive tests

| | Sensitivity | Specificity | Youden's index | AUC |
|---|---|---|---|---|
| **MMSE (*n* = 47)** | | | | |
| Paper | 0.55 | 0.75 | 0.30 | 0.65 |
| Digital | 0.91 | 0.75 | 0.66 | 0.82 |
| **CDT (*n* = 43)** | | | | |
| Paper | 0.64 | 0.47 | 0.11 | 0.45 |
| Digital | 0.45 | 0.50 | -0.05 | 0.65 |

Notes: AUC, Area under the curve; MMSE, Mini-Mental State Examination; CDT, Clock Drawing Test; SD, Standard Deviation

### Usability of digital cognitive tests

Of the 41 participants who completed the USE questionnaire, the digital tests received positive feedback on usefulness, ease of use, ease of learning, and satisfaction, with scores significantly higher than the median score ($P < 0.001$) (Table 4). Participants with lower education levels found it more difficult to learn how to use the tool ($P < 0.001$), and their overall USE mean score was lower ($P = 0.07$). However, half (55.26%) preferred the paper tools and only 23.68% preferred the digital tools.

The MMSE assessment took 6.21 ± 1.59 min for the paper-based version, significantly shorter than the 7.11 ± 2.64 min for the digital version ($P = 0.01$). Participants with lower education levels took longer to complete the assessments. No significant difference was found between the two versions of the CDT (Table 4).

### Regression analysis of digital cognitive test scores

The regression analysis revealed that participants' subjective attitudes of the digital test, including usefulness, ease of use, ease of learning, satisfaction, and preference, did not significantly impact the test scores (Table 5). In contrast, objective measures such as performance on the paper-based version ($\beta = 0.40$, $95\%CI = 0.07$ to $0.73$) and duration ($\beta = -0.62$, $95\%CI = -1.07$ to $-0.17$) were found to have a statistically significant effect on the digital test scores (Table 5).

## Discussion

This study utilized a randomized crossover design to examine the validity and usability of digital against paper-based cognitive tests in PHC settings. The findings demonstrated that, in terms of validity, eMMSE and eCDT showed moderate correlations with their paper-based counterparts, with digital cognitive tests achieving better AUC performance. Score discrepancies between digital and paper-based tests were primarily observed in the memory domain. Regarding usability, although older users provided positive feedback on digital tests, they still expressed a preference for paper-based tests, and digital tests required longer durations. Notably, older adults with lower education levels demonstrated weaker technological adaptability. Additionally, we found that digital cognitive tests scores were minimally influenced by users' subjective attitudes but were significantly associated with the objective factor of test duration.

Our study provides evidence supporting the validity of digital cognitive tests for screening in PHC settings. Both the eMMSE and the eCDT were moderately correlated with the paper versions suggesting that they have the potential to maintain screening power comparable

**Table 4** Usability of digital cognitive tests

| Variables | All sample | Education level | | | |
|---|---|---|---|---|---|
| | | Primary | Middle | High | *P* |
| **The USE questionnaire (*n* = 41)** | | | | | |
| Usefulness | 6.13 ± 0.85*** [a] | 5.91 ± 1.37 | 6.33 ± 0.56 | 6.01 ± 0.84 | 0.42 |
| Ease of use | 5.79 ± 1.09*** [a] | 5.21 ± 1.61 | 6.22 ± 0.70 | 5.6 ± 1.06 | 0.07 |
| Ease of learning | 5.49 ± 1.62*** [a] | 3.86 ± 1.85 | 6.35 ± 0.89 | 5.31 ± 1.59 | **0.001** |
| Satisfaction | 6.03 ± 0.96*** [a] | 5.59 ± 1.50 | 6.27 ± 0.51 | 5.97 ± 1.03 | 0.28 |
| Total average score | 5.90 ± 0.93*** [a] | 5.30 ± 1.46 | 6.28 ± 0.50 | 5.76 ± 0.93 | **0.04** |
| **Preference choice (*n* = 38)** | | | | | |
| Paper | 21 (55.26) | 3 (60.00) | 8 (50.00) | 10 (58.82) | 0.46 |
| Digital | 9 (23.68) | 1 (20.00) | 6 (37.50) | 2 (11.77) | |
| No preference | 8 (21.05) | 1 (20.00) | 2 (12.50) | 5 (29.41) | |
| **Duration (*n* = 47)** | | | | | |
| MMSE | 6.21 ± 1.59 | 7.89 ± 1.54 | 5.84 ± 1.07 | 5.79 ± 1.58 | **0.001** |
| eMMSE | 7.11 ± 2.64** [b] | 9.22 ± 3.56 | 7.42 ± 2.41** [b] | 5.79 ± 1.47 | **0.003** |
| CDT | 4.38 ± 2.57 | 5.33 ± 2.69 | 2.89 ± 2.11 | 5.42 ± 2.32 | 0.003 |
| eCDT | 4.57 ± 2.71 | 5.33 ± 2.96 | 3.68 ± 1.83 | 5.11 ± 3.18 | 0.18 |

Notes: Values are shown as Mean ± SD or *n* (%). USE, Usefulness, Satisfaction, and Ease of Use; MMSE, Mini-Mental State Examination; CDT, Clock Drawing Test; eMMSE, electronic versions of the Mini-Mental State Examination; eCDT, electronic versions of Clock Drawing Test; MCI, Mild cognitive impairment; NC, Normal cognition

[a] The one-sample Wilcoxon signed rank test compared the mean scores of the USE to the median of the Likert scale 4

[b] Paired t-tests were used to compare the duration of the paper and electronic cognitive tests

**$P < 0.01$, ***$P < 0.001$

**Table 5** Regression model for eMMSE and eCDT

| | eMMSE | | eCDT | |
|---|---|---|---|---|
| | **Coefficient** | **95% CI** | **Coefficient** | **95% CI** |
| **Age** | 0.02 | (-0.12, 0.17) | 0.10 | (-0.18, 0.38) |
| **Sex** | 1.38 | ( -0.19, 2.94) | 1.46 | (-0.47, 3.39) |
| **Education level** | | | | |
| Primary | Ref. | | | |
| Middle | 1.24 | ( -1.52, 4.00) | 0.18 | (-2.73, 3.08) |
| High | 1.46 | ( -1.25, 4.17) | 1.39 | (-0.53, 3.31) |
| **Paper version score** | 0.40 | **(0.07, 0.73)** | 0.71 | **(0.45, 0.97)** |
| **The USE questionnaire** | | | | |
| Usefulness | -0.55 | ( -1.70, 0.60) | -0.88 | (-1.98, 0.22) |
| Ease of use | -0.69 | ( -2.13, 0.74) | 0.30 | (-0.98, 1.59) |
| Ease of learning | 0.11 | ( -0.71, 0.93) | 0.48 | (-0.36, 1.33) |
| Satisfaction | 0.92 | ( -0.49, 2.34) | -0.01 | (-1.42, 1.41) |
| **Preferences** | | | | |
| Paper | Ref. | | | |
| Electronic | 0.98 | ( -1.39, 3.36) | -1.11 | (-3.26, 1.03) |
| No preference | -0.06 | ( -2.19, 2.07) | -1.57 | (-4.60, 1.45) |
| **Duration** | -0.62 | **(-1.07, -0.17)** | 0.27 | (-0.34, 0.87) |

Notes: USE, Usefulness, Satisfaction, and Ease of Use; eMMSE, electronic versions of the Mini-Mental State Examination; eCDT, electronic versions of Clock Drawing Test. Interaction terms for education level*duration and education level * USE were tested and found to be non-significant

to paper-based cognitive tests, in line with prior findings for eMoCA [9, 15]. Our study also showed that the eMMSE demonstrates higher sensitivity to distinguish between normal individuals and those with MCI compared to the MMSE. Further analysis by domain of MMSE indicates that this difference is mainly attributed by memory domain, namely the immediate recall and delayed recall tasks. Similar findings had been reported previously regarding SLUMS [24] and MoCA [15]. As these tasks were pre-recorded and automatically broadcasted during the digital cognitive tests, it may incur extra cognitive load on older adults as they need to pay attention to these monotonous records without personal interaction [15, 24]. Particularly in settings where external distraction exists, older adults are easily lost to follow the tests. Additionally, older adults may find it challenging to understand synthetic speech outputs from digital devices, which can be exacerbated by age-related changes in hearing [15]. This extra cognitive efforts in memory domain due to its implementation features, may thus be more sensitive to distinguish cognitive normal versus MCI participants.

This study indicates that digital cognitive testing holds great potential for practical application in PHC settings, but achieving this goal requires overcoming challenges related to user preferences and technological adaptability. Overall, although older adults have a positive attitude towards the functionality and operability of digital tests, believing they met their needs and expectations, many still preferred paper-based tests in practice. This highlights a disconnect between their subjective perceptions and actual preferences. Open-ended responses further supported this observation, revealing a common sense of unfamiliarity with digital devices and their interfaces among older users. These challenges were particularly pronounced in older adults with lower educational levels. They faced greater difficulties in adapting to digital testing technology, as evidenced by their subjective perception that digital tests were harder to learn and their more negative attitudes toward system usability. These findings are consistent with recent studies [14, 25, 26]. Additionally, we observed that older adults with lower educational levels were objectively more likely to take longer to complete the tests, primarily due to unfamiliarity with digital devices and a lack of confidence [27]. These results further highlight the significant challenges this population faces in terms of technological adaptability.

Although older participants had varying subjective attitudes and preferences toward the digital test format, these subjective factors did not have a significant impact on their performance. This finding is consistent with previous research on the eMoCA [15], which also found no correlation between user preferences and test scores. Furthermore, our study revealed that objective measures, such as test duration, serve as more sensitive indicators, better reflecting the practical challenges older adults face when performing digital tests. After controlling for cognitive level, education, age, sex, USE and preference, we observed that longer eMMSE durations were associated with lower scores. This suggests that other factors such as adaptability to digital devices [15, 28, 29] or anxiety and tension regarding the unfamiliar test format [30] may contribute to longer test durations and poorer performance. These findings emphasize the need to consider

the impact of the digital test format when designing and implementing digital assessments. Addressing operational and psychological challenges faced by older adults through strategies like enhanced guidance, optimized interfaces, and tailored procedures is essential to reduce anxiety, improve user experience, and ensure reliable results [31, 32].

### Limitations of the study and future work

This study has several limitations. Firstly, the relatively small sample size may restrict the generalizability of the findings, particularly in populations with diverse educational or digital literacy backgrounds. Secondly, usability assessments relied on participants' self-reports data, which may be subject to social desirability. Third, due to questionnaire and technical limitations, the eMMSE and eCDT did not feature automated scoring or intelligent algorithms. Test administrators were still required to manually score the tests, which may have introduced some degree of subjective biases. Fourth, although a two-week washout period was implemented to mitigate potential carryover effects in the crossover design, the possibility of residual period effects cannot be entirely excluded. Fifth, While MMSE and CDT are commonly employed as reference standards for evaluating criterion validity in cognitive assessments, both tools have known psychometric limitations. Their limited sensitivity, particularly in detecting early cognitive impairment, may introduce reference standard bias and affect the accuracy of validity estimates. Finally, neurological diagnoses were conducted only for screen-positive participants and approximately 10% of screen-negative individuals, which may have introduced verification bias and potentially inflated the estimates of sensitivity and specificity.

Future research should consider larger and more diverse samples to enhance external validity. Additionally, the integration of advanced digital technologies—such as artificial intelligence and machine learning—should be prioritized to develop intelligent, automated scoring systems. Incorporating features like real-time response tracking, reaction time analysis, and eye movement monitoring could enhance the objectivity, sensitivity, and diagnostic accuracy of digital assessments.

## Conclusion

In conclusion, digital cognitive tests are valid and feasible in PHC settings, demonstrating outstanding performance in both validity and users' subjective attitudes toward usability. However, several challenges still need to be overcome during the promotion process, particularly regarding users' preference for paper-based tests and the technological adaptability issues faced by individuals with lower education levels. By enhancing training and optimizing technical support, digital cognitive tests are expected to become efficient and scalable cognitive screening tools in PHC, especially in resource-limited settings where they hold significant application potential.

#### Abbreviations

| | |
|---|---|
| AUC | Area under the curve |
| MMSE | Mini-Mental State Examination |
| eMMSE | electronic Mini-Mental State Examination |
| CDT | Clock Drawing task |
| eCDT | electronic Clock Drawing Task |
| MCI | Mild cognitive impairment |
| PHC | Primary health care |
| USE | Usefulness, Satisfaction, Ease of use |
| MoCA | Montreal Cognitive Assessment |
| eMoCA | electronic Montreal Cognitive Assessment |
| SLUMS | Saint Louis University Mental Status Examination |

#### Acknowledgements

We thank all the participants for their participation in this study. We appreciate the professional training and diagnostic expertise provided by Zhuohui Huang (Guangdong Mental Health Center) and Ruoyan Huang (Guangzhou Mental Health Center). We also thank Daoyu Wang and Hao Zeng (Shanghai Zhuodao Medical Technology Co., Ltd.) for providing equipment and technical support.

#### Author contributions

MW conceived and designed the study, collected participants' basic information, analyzed and interpreted the data, assisted in preparing tables and figures, and drafted the initial manuscript. JLF analyzed and interpreted the data, assisted in preparing tables and figures, and co-drafted the initial manuscript. RNS, YJZ, and SYZ assisted in the study design, collected participants' basic information, and participated in on-site quality control. FJY, XYZ, and YXY recruited participants, determined participant eligibility, coordinated the study, and conducted MMSE, eMMSE, CDT, and eCDT assessments. JL and NG conceived and designed the study, supervised data analysis and interpretation, and revised the manuscript. All authors reviewed and approved the final version for submission. Min Wu and Jialin Feng contributed equally to this work.

#### Funding

This study was supported by the National Science Foundation of China/ the Economic and Social Research Council, UK Research and Innovation joint call: Understanding and Addressing Health and Social Challenges for Ageing in the UK and China, UK-China Health and Social Challenges Ageing Project (UKCHASCAP): present and future burden of dementia, and policy responses (Grant number 72061137003, ES/T014377/1), but they had no role in the design of the study, data collection, analysis, interpretation of the data, or in writing the manuscript.

#### Data availability

The data supporting the findings of this study are not publicly available due to privacy and confidentiality concerns for the participants. Data can be made available upon reasonable request from the corresponding author, subject to ethical and legal restrictions.

## Declarations

#### Ethics approval and consent to participate

This study meets the institutional requirements for human trials and has been approved by the Biomedical Research Ethics Committee of the School of Public Health, Sun Yat-sen University, China (ethics approval number: 2023 – 129). It was conducted in accordance with the ethical standards set forth in the 1964 Declaration of Helsinki and its subsequent amendments. All subjects provided written informed consent voluntarily.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 30 December 2024 / Accepted: 4 June 2025
Published online: 11 June 2025

**References**

1. Jia J, et al. The cost of alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement. 2018;14(4):483–91.
2. Lissek V, Suchan B. Preventing dementia? Interventional approaches in mild cognitive impairment. Neurosci Biobehav Rev. 2021;122:143–64.
3. Gong N, et al. Exploring barriers to dementia screening and management services by general practitioners in china: a qualitative study using the COM-B model. BMC Geriatr. 2023;23(1):55.
4. Belleville S, LaPlume AA, Purkart R. Web-based cognitive assessment in older adults: where do we stand? Curr Opin Neurol. 2023;36(5):491–7.
5. Tsoy E, Zygouris S, Possin KL. Current state of Self-Administered brief computerized cognitive assessments for detection of cognitive disorders in older adults: A systematic review. J Prev Alzheimers Dis. 2021;8(3):267–76.
6. Cubillos C, Rienzo A. Digital cognitive assessment tests for older adults: systematic literature review. JMIR Ment Health. 2023;10:e47487.
7. Chan JYC et al. Electronic cognitive screen technology for screening older adults with dementia and mild cognitive impairment in a community setting: development and validation study. 2020;22(12):e17332.
8. Wong A, et al. Computerized cognitive screen (CoCoSc): A Self-Administered computerized test for screening for cognitive impairment in community social centers. J Alzheimers Dis. 2017;59:1299–306.
9. Yu K, et al. Development of a computerized tool for the Chinese version of the Montreal cognitive assessment for screening mild cognitive impairment. Int Psychogeriatr. 2015;27(2):213–9.
10. Piccinin AM, et al. Coordinated analysis of age, sex, and education effects on change in MMSE scores. J Gerontol B Psychol Sci Soc Sci. 2013;68(3):374–90.
11. de Noronha Í, Barreto FCSDS, Ortiz KZ. The influence of education on performance of adults on the clock drawing test. Dement Neuropsychol. 2018;12(1):61–7.
12. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc. 1992;40(9):922–35.
13. Valladares-Rodriguez S, et al. Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. PeerJ. 2017;5:e3508.
14. Lunardini F, et al. Supervised digital neuropsychological tests for cognitive decline in older adults: usability and clinical validity study. JMIR Mhealth Uhealth. 2020;8(9):e17963.
15. Wallace SE, et al. A comparison of electronic and paper versions of the Montreal cognitive assessment. Alzheimer Dis Assoc Disord. 2019;33(3):272–8.
16. WHO. The International Classification of Diseases, 11th revision. 2019; Available from: https://icd.who.int/en
17. Petersen RC, et al. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol. 1999;56(3):303–8.
18. Scharre DW, et al. Digitally translated Self-Administered gerocognitive examination (eSAGE): relationship with its validated paper version, neuropsychological evaluations, and clinical assessments. Alzheimers Res Ther. 2017;9(1):44.
19. Folstein MF, Folstein SE, and McHugh PR, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189–98.
20. Wang B, et al. The clinical characteristics of 2,789 consecutive patients in a memory clinic in China. J Clin Neurosci. 2011;18(11):1473–7.
21. Shon JM, et al. Functional neuroanatomical correlates of the executive clock drawing task (CLOX) performance in alzheimer's disease: a FDG-PET study. Neuroscience. 2013;246:271–80.
22. Royall DR, Cordes JA, Polk M. CLOX: an executive clock drawing task. J Neurol Neurosurg Psychiatry. 1998;64(5):588–94.
23. Lund AM. Measuring usability with the USE questionnaire. STC Usability SIG Newsl. 2001;8(2).
24. Ruggeri K, et al. Are we there yet?? Exploring the impact of translating cognitive tests for dementia using mobile technology in an aging population. Front Aging Neurosci. 2016;8:21.
25. Dwolatzky T, et al. Validity of a short computerized assessment battery for moderate cognitive impairment and dementia. Int Psychogeriatr. 2010;22(5):795–803.
26. Hargittai E, Dobransky KJCJoC. Old dogs, new clicks: digital inequality in internet skills and uses among older adults. 2017;42(2):195–212.
27. Snowdon A, et al. Comparison of an electronic and Paper-based Montreal cognitive assessment tool. Alzheimer Dis Assoc Disord. 2015;29(4):325–9.
28. Raymond PD, et al. Test-retest norms and reliable change indices for the microcog battery in a healthy community population over 50 years of age. Clin Neuropsychol. 2006;20(2):261–70.
29. Wallace SE, et al. Validation of the standardized touchscreen assessment of cognition with neurotypical adults. NeuroRehabilitation. 2017;40(3):411–20.
30. Eraslan Boz H, et al. A new tool to assess amnestic mild cognitive impairment in Turkish older adults: virtual supermarket (VSM). Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2020;27(5):639–53.
31. Friemel T. N.J.N.m. And society, The digital divide has grown old: determinants of a digital divide among seniors. 2016;18(2):313–31.
32. Van Deursen AJ, J.A.J.N.m., Van Dijk. and society, The digital divide shifts to differences in usage. 2014;16(3):507–526.

## Publisher's note